DOCUMENT RESUME

ED 078 010 TM 002 822

AUTHOR Yoshida, Roland K.; And Others
TITLE A Psychometric Approach to Piaget: Some Theoretical and Methodological Implications.
PUB DATE [73]
NOTE 18p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Behavior Rating Scales; Child Development; *Cognitive Development; Cognitive Tests; Individual Tests; Learning Theories; Literature Reviews; Psychometrics; *Test Construction; *Testing Problems
IDENTIFIERS *Piaget (Jean)

ABSTRACT

The sequence of development as hypothesized by Piaget has been confirmed for various groups of children by cross-sectional and longitudinal studies. Such consistency has increased interest in constructing scales based on Piaget's theory according to psychometric principles. Three major problems are discussed which present definite obstacles in the development of a general scale: (1) A Piagetian test, encompassing all stages and operations within individual states, would be so broad in scope that thorough testing may be precluded; (2) Piaget's clinical method of questioning may not be amenable to standardization, resulting in unreliability from a psychometric standpoint; and (3) Piagetian demonstrations have been presented in many different forms and at various degrees of complexity; such diversity necessitates an imposing task of selecting a set of items most representative of the sequency of development. Considering these overriding problems, it is concluded that one scale representing Piagetian development in general may be an impossibility. (Author/KM)

ED 078010

TM 002 822

Abstract

The sequence of development as hypothesized by Piaget has been confirmed for various groups of children by cross-sectional and longitudinal studies. Such consistency has increased interest in constructing scales based on Piaget's theory according to psychometric principles. Three major problems were discussed which present definite obstacles in the development of a general scale. (1) A Piagetian test, encompassing all stages and operations within individual stages, would be so broad in scope that thorough testing may be precluded. (2) Piaget's clinical method of questioning may not be ameanable to standarization, hence, unreliability from a psychometric standpoint. (3) Piagetian demonstrations have been presented in many different forms and at various degress of complexity; such diversity necessitates an imposing task of selecting a set of items most representative of the sequence of development. Considering these overriding problems, the authors concluded that one scale representing Piagetian development in general may be an impossibility.

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

FILMED FROM BEST AVAILABLE COPY

A Psychometric Approach to Piaget: Some Thecretical and Methodological Implications

Roland K. Yoshida, C. Edward Meyers
University of Southern California
and Russel E. Orpet
California State University, Long Beach

The construction of scales based upon Piaget's theory has gained considerable support from various groups of professionals (Pinard & Laurendeau, 1964; Sullivan, 1967; Woodward, 1963). Piaget himself once expressed such an interest (Piaget & Inhelder, 1947) though he has left the task of test construction to others. The central theme underlying this movement is the argument that currently used intelligence tests have fulfilled statistical but not logical criteria for a "good" test. Moreover, Binet and Wechsler type tests have been challenged on the following grounds: a) they measure outcomes rather than processes, b) they are biased in sampled content and in norming with only white middle class children (Williams, 1971), and c) their difficulty levels are established on statistical bases, not inherent difficulty. The marketed intelligence tests are merely attempting to discriminate among children at various ages, using items chosen for their ability to do so. A more acceptable approach to

mental testing is one whose items are selected on the basis of a genuine theory of cognitive development. Piaget's theory offers just such a possibility.

In an address presented to this conference last year, Meyers (1972) posed the following question: Can Piaget's theory provide a better psychometry? Although some optimism was suggested for limited scales, a set of problems was forwarded which could ultimately question the validity of scales based on Piaget's theory. It is the purpose of this paper to discuss those theoretical and methodological questions in more detail with the intent of cautioning interested professionals to the issues of applying psychometric principles to Piagetian theory in the hopes of producing useful and valid scales.

The first step in constructing scales based on a theory is to determine whether the theory is a valid one; that is, internally consistent and amenable to testing. Innumerable studies have replicated many of Piaget's stages and demonstrations showing them to be in true ordinal relationship to one another for various groups of children, including the developmentally disabled (Inhelder, 1968; Reiss, 1968; Yoshida, 1973). A wide range of ages mark the onset of different operational behaviors but the order of their appearance remains invariant. The stages of tasks such as conservation items of the Goldschmid-

Bentler (1968) variety form in most cases a Guttman scale which is sensitive to the hierarchical sequence of items in which the most difficult item succeeds the easier ones and so on. In short, we have a tentative validation of the theory.

The test development work, like most Piagetian research, thus far has almost exclusively employed cross-sectional designs. The usual procedure selects subjects according to age, then noting whether their responses form a Guttman scale. This procedure has the usual drawbacks of a cross-sectional study, namely that we are not completely sure that individual children follow a particular pattern of development over the long-term, particularly whether they demonstrate an order that deviates from the hypothesized sequence of development. In other words, order found by cross-sectional techniques will always be tentative. A longitudinal study must be conducted to insure the validity of the hypothesized order.

We have knowledge of two longitudinal efforts. One project by Almy and others (1966) confirmed that conservation of number precedes that of continuous quantity or liquid conservation as was found in replication studies. An ongoing project by Stephens at Temple University using normal and mentally retarded subjects (Stephens, McLaughlin, & Mahaney, 1971) has yet to

publish complete findings on the order of stage development. Their preliminary data appear to indicate correspondence with Piaget's findings. Though narrow in domains examined, longitudinal studies have resulted in a qualified validation of Piaget's stages.

Finding general validation does not give us sufficient reason to accept Piaget as an alternative to the current measures. We must also ascertain whether Piagetian tasks measure a construct different from those sampled by present tests or do better in the measurement of what is intended to be measured. Kaufman (1971) factor analyzed a test battery consisting of the Lorge-Thorndike Intelligence test, Gesell School Readiness test and a Piagetian battery composed of some conservation, class inclusion, logic and geometric problems to a group of elementary school children. Each test loaded on separate factors, demonstrating the relative independence of Piagetian tasks from the content of other psychometric instruments. Meyers and Orpet (1971) found with 5½ year olds that their selected conservation tasks did not load on a single factor but were factorially complex. Although some of the tasks loaded on factors containing some WPPSI, Raven and ITPA subtests, there was a high degree of specificity for the Piagetian tasks in relationship to the various psychometric instruments. Finally, correlations of Piagetian tasks with

MA, IQ, and the subtests of the WISC are generally low, rarely exceeding .50 (Dodwell, 1961; Elkind, 1961; Goldschmid, 1967); the highest r's, +.52-+.62, were obtained by Dudek, Lester, Goldberg, and Dyer (1969). Even with an MA range of 2½ to 8 years in trainable mentally retarded subjects, Yoshida (1973) found Kendall tau coefficients of only +.32 or lower between Binet MA and Guttman scale scores of Piaget-Inhelder haptic perception tasks. Thus, Piagetian tasks may tap aspects of a construct of intelligence which are quite different from those sampled by the WISC and other psychometric instruments.

Thus far we have developed two conclusions. One, the Piagetian scales have been shown to follow the general sequence hypothesized by Piaget and his co-workers over different test situations and subjects. Two, they are substantially independent of the constructs measured by traditional instruments. Considering these findings, we have strengthened the rationale for building standardized Piagetian scales which may be used in conjunction with or eventually replace the Binet and its derivatives.

Assuming that we accept Piagetian theory as our standard of development, what are the problems encountered in constructing a useful test or battery of tests? First of all, Piaget's theory spans a large section of time analyzing levels of development in various concepts.

We have to contend with time and content dimensions which may not be readily reduced into one scale. Secondly, we are well aware of the published protocols of Piaget and his colleagues in which the investigator interacts with his subject. The obvious difficulty is extracting those aspects of Piaget's questioning technique which are modifiable for easy wording and in a more important sense, standardization. Not to be ignored are the many ways within demonstrations such as class inclusion which were varied by Piaget to determine the qualitative aspects of growth. What variations in materials would make for the most valid diagnosis have yet to be determined. As a result, we have an enormous albeit rich corpus of materials and methods from which to select our test items. If we reduce this body for the sake of efficiency and standardization as a psychometric approach would necessitate, would we be compromising the validity of our final product? What follows is a discussion centering on the issues of inclusiveness of tests, the method of questioning and the complexity of test materials based on Piaget's theory.

Let us begin with the issue of inclusiveness. Piaget's mental development theory encompasses a fairly large chronological age range of a typical child's life. Certain classes of behaviors and mental structures have

been hypothesized for neonates through adolescents. Although the development of infant scales by Uzgiris and Hunt (1968) and Escalona and Corman (1967) are of extreme importance, they will not be typically employed in the school setting. We need to concentrate on processes beginning at two years and culminating with the higher levels of formal operational thinking. What we are talking about is a test which samples behavior from 2 to 18 years.

The most extensively reported prototype of such a general scale is the one devised by Pinard and Laurendeau (1964) comprising of 57 subtests beginning with items from the preoperational stage and extending to the formal operations stage. Twenty-five subtests were directly taken from Piaget's work in the areas of causality, time, movement, speed, relations, number, space and conservation. The attempt by the constructors to devise a comprehensive scale across most stages and content domains reveals the extreme caution taken to insure proper placement of the examinee on a general scale of development. Altogether the test required 10 hours on the average to administer and was divided into four to six sessions depending on the child's age. Use of the clinical method may explain in part the prohibitive amount of time taken to adminster the battery. Nevertheless, the sheer number of tasks contributed to most of the time involved.

The above effort exemplifies the possible breadth of a Piagetian test battery. Our goal is to assess a child in relation to an operational level, not limiting ourselves to a certain type of demonstration. Published or quasi-formal Piagetian test instruments are quite limited in scope. For example, the Goldschmid-Bentler Conservation Assessment Kit taps only conservation of some of the more popular tasks such as substance, weight, liquid, number and so on. Processes such as seriation and transitivity and other tasks drawn from the concrete operational stage are not sampled. Accordingly, the test inadequately assesses children who have reached this level of thinking and the generality of results are restricted to conservation only.

What we are saying is that conservation is only a partial sampling of the concrete operational stage. By limiting the test to only conservation as an example, we may make judgments to specific content areas without more precisely pinpointing development of a more general nature. That is, can children who correctly solve conservation problems also respond operationally to class inclusion problems which have been hypothesized to be at the same level of development?

In general, Piagetian demonstrations load on separate factors and produce low intercorrelations among themselves.

Lunzar (1970) sampling the conservation, classification, logic and spatial domains found that those Piagetian tasks yielded four factors. The largest factor in terms of variance accounted for loaded with most of the items from the battery. However, the other three factors, classification, spatial and what Lunzar termed a verbal factor comprised 40% of the total variance, demonstrating the diversity of the abilities sampled. Kaufman also found separate factors from his battery of Piagetian tasks. Another innovative experimenter, Tuddenham of the University of California, Berkeley (1971) found low intercorrelations among his tasks. Unreliability was ruled out as a possible explanation for such a diffuse correlation matrix. Similarly, Goldschmid and MacFarlane (1968) found correlations on the order of +.25 between scores from Form A of the Conservation Assessment Kit with tests of probability, seriation, classification and perspective. With the exception of the latter investigators, specificity of tasks was forwarded as the probable cause for the above results.

The four studies indicate that although we are dealing with a single course of development, specific operations are manifested noncorrespondingly across the different content areas. In order to accurately and validly assess growth of the total operational system, we must sample a wide variety of tasks.

Our second point of controversy concerns Piaget's use of the clinical method to probe the child's ability to understand a task. Briefly, Piagetian tasks are presented in the following manner:

a) Language check is made before commencing the demonstration or judgment to this effect during testing.

b) Agreement to the properties of the tasks such as in a classification task that all the beads are wooden, their colors and other properties as size are the same.

c) Establishment of equivalence if necessary as in a typical conservation problem as liquid in which water is at the same level in all beakers.

d) Perceptual transformation of the objects or some change in the presented situation.

e) Judgment question.

f) Explanation questions.

The first five steps are dispensed with rather quickly. However, the child must explain his judgment answer. A Piagetian process does not merely ask a question to a given problem, recording the responses as right or wrong. It probes into how the child reached his conclusion or judgment by challenging and countersuggesting to determine whether the answer was contrived given this response "my teacher told me so," perceptually oriented as "they both look the same," or a true mastery of the concept.

What we have here is the desire to increase

confidence in assessment. The distinction is often made between appraising what a child does given standard objective questions in contrast to what he can be brought to do with the appropriate probing of the clinical method. This procedure unsheaths a two-edged sword. On the one hand, we could disuade a child who is unsure of his answer. We could uncover the fact that his schema is not completely intact. On the other hand, we might tease out of a youngster an acceptable conceptual response after he unwittingly gave a nonoperational response as "we found out in science." Quite possibly he did but if he can adequately explain the concept we would change our conclusion to one of the subject thinking at a higher operational level.

The inclusion of the clinical method in test instruments has become an important issue. If the clinical method is drastically altered for the sake of convenience and efficiency or is eliminated entirely as some writers have suggested (Brainerd, 1973), are we really removing most of the characteristics of a valid Piagetian exploration as Inhelder, Bovet, Sinclair and Smock (1966) contend?

Quasi-clinical methods have been devised which standardize questioning but allow individual discretion in deviating from the set pattern (Lunzar, 1970; Tuddenham,

1971). Tuddenham in particular states that the clinical method a la Piaget was ideal for discovering qualitative differences in a child's thought and for formulating a theory of development. However, such a technique interferes with the presentation of materials under identical conditions which is the psychometric approach. Tuddenham argues that the groundwork for substantiating the theory has been completed and that for test purposes liberties may be taken with the method clinique. His modified version of the clinical method is in the spirit of Piagetian inquiry. He allows the child to explain his judgment. When ambiguity as to the nature of responses occur, standard questions are then administered to fit the individual case. Rigorous investigation must be conducted between results using the classical method and any of the alternative procedures before any one of them is accepted. We could thus have the best of both worlds.

Our third problem is that of task complexity of materials.which may have specific effects on results. Feigenbaum (1963) varied the number of beads and the perceptual disparity between containers in a discontinuous quantity problem. In his experiment, 12 beads were presented in one condition and 24 in another. The physical size difference between the original glass containers and the one the beads were to be poured into

was reduced to produce less perceptual distortion. The findings were somewhat mixed. The increased number of beads resulted in fewer conserving responses for children who were concluded to have incomplete operational structures. The size of the containers had no effect on the frequency of correct responses for the entire group of subjects. Goldschmid (1967) also reported differences in correct judgments with conservation of substance, continuous and discontinuous quantity when complexity of materials was manipulated. The above evidence suggests that constructing items may not be a simple undertaking. Attention must be given the variables affecting the item difficulty for each Piagetian task.

This paper has reviewed the basic rationale for developing tests based upon Piaget's mental theory. We have discussed the problems of inclusiveness, method of interaction with the examinee and the complexity of tasks selected for inclusion on such a test. No doubt, the problems are formidable and possibly insurrmountable. Test constructors are confronted with the predicament of preserving the consistency between theory and a test instrument while at the same time fulfilling the requirements of standardization, brevity and efficiency. These two approaches may not be compatible as one would wish. And the compromising of the two may not result in an

acceptable measure to either Piagetian purists or those who desire a condensed testing instrument.

A parsimonious scale then by our view is almost out of the question. The restrictions due to time and the finding of separate content domains within stages preclude such a culmination. Rather limited scales testing individual functions will be more within reason. However, even that scenario may not be completed until issues such as "the exploration method," item difficulty of tasks and the order of item presentation are resolved. The path before us is a formidable one; one that demands patience, suggestion and countersuggestion to achieve our goal.

References

Almy, M., Chittenden, E., and Miller, P. Young children's thinking: Studies of some aspects of Piaget's theory. New York: Teacher's College Press, 1966.

Brainerd, C. J. Judgments and explanations as criteria for the presence of cognitive structures. Psychological Bulletin, 1973, 79, 172-179.

Dodwell, P. C. Children's understanding of number concepts: Characteristics of an individual and of a group test. Canadian Journal of Psychology, 1961, 15, 29-36.

Dudek, S. Z., Lester, E. P., Goldberg, J. S., and Dyer, G. B. Relationship of Piaget measures to standard intelligence and motor scales. Perceptual and Motor Skills, 1969, 28, 351-362.

Elkind, D. The development of quantitative thinking: A systematic replication of Piaget's studies. Journal of Genetic Psychology, 1961, 98, 37-46.

Escalona, S. K., and Corman, H. H. The validation of Piaget's hypotheses concerning the development of sensori-motor intelligence: Methodological issues. Paper presented at a meeting of the Society for Research in Child Development, New York, 1967.

Feigenbaum, K. D. Task complexity and IQ as variables in Piaget's problem of conservation. Child Development, 1963, 34, 423-432.

Goldschmid, M.L.Different types of conservation and nonconservation and their relation to age, sex, IQ, MA and vocabulary. Child Development, 1967, 38, 1229-1246.

Goldschmid, M. L., and Bentler, P.M. Manual: Concept Assessment Kit- Conservation. San Diego, California: Educational and Industrial Testing Service, 1968.

Goldschmid, M. L., and MacFarlane, B. The assessment of six Piagetian concepts in the same subjects: Classification, conservation, perspective, probability, seriation and transitivity. Unpublished paper, McGill University, 1968.

Inhelder, B. The diagnosis of reasoning in the mentally retarded. New York: John Day, 1968.

Inhelder, B., Bovet, M., Sinclair, H., and Smock, C. D. On cognitive development. American Psychologist, 1966, 21, 160-164.

Kaufman, A. S. Piaget and Gesell: A psychometric analysis of tests built from their tasks. Child Development, 1971, 42, 1341-1360.

Lunzar, E. A. Construction of a standardised battery of Piagetian tests to assess the development of effective intelligence. Research in Education, 1970, 3, 53-72.

Meyers, C. E. Can Piaget's theory provide a better psychometry? Proceedings, Second Annual UAP Conference: Piagetian Theory and the Helping Professions, 1972, 5-10.

Meyers, C. E., and Orpet, R. E. Ability factor location of some Piagetian tasks at 5½ years. Proceedings, 79th Annual Convention, American Psychological Association, 1971, 199-200.

Piaget, J., and Inhelder, B. Diagnosis of mental operation and theory of intelligence. American Journal of Mental Deficiency, 1947, 51, 401-406.

Pinard, A., and Laurendeau, M. A scale of mental development based on the theory of Piaget: A description of a project. Journal of Research in Science Teaching, 1964, 2, 253-260.

Reiss, P. Implication of Piaget's developmental psychology for mental retardation. American Journal of Mental Deficiency, 1967, 72, 361-369.

Stephens, B., McLaughlin, J. A., and Mahaney, E. J. Ages at which Piagetian concepts are achieved. Proceedings, 79th Annual Convention, American Psychological Association, 1971, 203-204.

Sullivan, E. V. Piaget and the school curriculum- A critical appraisal. The Ontario Institute for Studies in Education, No. 2, 1967.

Tuddenham, R. D. Theoretical regularities and individual idiosyncrasies. In D. R. Green, M. P. Ford, and G. B. Flamer (Eds.), Measurement and Piaget. New York: McGraw-Hill, 1971, 64-75.

Uzgiris, I. C., and Hunt, J. McV. Ordinal scales of infant psychological development: Information concerning six demonstration films. Mimeographed paper, Psychological Developmental Laboratories, University of Illinois, 1968.

Williams, R. L. Abuses and misuses in testing black children. The Counseling Psychologist, 1971, 2, 62-73.

Woodward, M. The application of Piaget's theory to research in mental deficiency. In N. R. Ellis (Ed.), Handbook of mental deficiency. New York: McGraw-Hill, 1963. Pp. 297-324.

Yoshida, R. K. A Guttman scalogram analysis of haptic perception for trainable mentally retarded children. American Journal of Mental Deficiency, 1973, 77, 439-444.